

STUDENT RATING OF TEACHERS

A Thesis

Presented to  
the School of Education  
Drake University

In Partial Fulfillment  
of the Requirements for the Degree  
Educational Specialist

by Jerry M. Deegan

October 1993

LOCKER  
1993  
.D36  
C.2

ACC 4553

STUDENT RATING OF INSTRUCTION


Jerry M. Deegan


October 1993

Approved by Committee:

  
Edward R. Ducharme

  
Mary K. Ducharme

  
Richard L. Schwab

  
Richard L. Schwab  
Dean of the School of Education

## STUDENT RATING OF TEACHERS

An abstract of a Thesis by

Jerry M. Deegan

October 1993

Drake University

Advisor: Edward R. Ducharme

The problem. In 1987, Central High School, located in a large midwestern city, instituted a system allowing student input to be a part of the teacher evaluation process. After five years of data collection, it was decided there would be value in studying the results of the ratings in relation to specific components of effective instruction.

Procedures. Student ratings of teachers from 1990-1992 were used for the study. From the instrument, 10 items were selected that were directly related to five elements of effective instruction. The five teachers whose results showed the greatest increases were chosen for interviews.

Findings. Common themes were found among the teachers' perceptions of the practice and results of student ratings. Teachers valued the student input and shared ideas as to how to improve the process.

Conclusions. Student ratings are a valuable part of the teacher evaluation process. Teachers generally welcome the feedback and value what students have to say. Teachers want to have input into the process and creation of the instrument.

Recommendations. Student rating of teachers should be included in the teacher evaluation process. Students should be asked to do ratings for all classes. Teachers should be given opportunities for input in the creation of the instrument and the time of administration. The instrument should include items rating effective components of instruction and course content. Students should be informed and included in the process. The data gathered should be available to teachers in a timely fashion and tracked over time.

## TABLE OF CONTENTS

	Page
Chapter	
1. INTRODUCTION .....	1
2. REVIEW OF THE LITERATURE .....	16
3. PROCEDURES .....	30
4. ANALYSIS AND INTERPRETATION .....	38
5. SUMMARY AND CONCLUSIONS .....	51
REFERENCES .....	55
APPENDIXES	
A. RATING INSTRUMENT USED BY STUDENTS .....	58
B. INDIVIDUAL TEACHER RESULTS .....	61
C. INTERVIEW QUESTIONNAIRE .....	71

## Chapter 1

### INTRODUCTION

To describe the roots of this study, I must go back to the spring of 1987. At that time, I was a school counselor at a large metropolitan high school working on completion of certification in secondary administration. I asked the principal at Central High School if I could do my internship with him. He agreed and we discussed the possible goals and activities of my internship. Among the interests he and I both had was to study the current teacher evaluation system at Central High and to report the findings. With his approval, I decided to form a committee of faculty to assist me in the process. I sought volunteers and ended up with three teachers and myself as the committee. One of the teachers had been a teacher at Central High for 32 years with over 38 years in education. Another of the teachers had 10 years of teaching experience, 9 of which were at Central High. And the third teacher had 21 years of teaching experience, 12 of which were at Central High. In selecting committee members, I sought a cross section of the faculty and teachers who were generally regarded by a majority of the staff as competent teachers. One of these teachers was a Teacher of the Year at Central High, the second was department chair in business for 15 years,

and the third was a teacher who assisted with volunteer activities across all the grade levels.

The committee met during the semester and shared information from research regarding teacher evaluation programs. Ten schools in the surrounding area as well as Catholic schools of similar size in the state were contacted and asked for copies of their teacher evaluation policies and procedures. The committee used the information from research and schools to construct three surveys. These surveys were given to the staff over the second semester of the year. After each, the results were shared with the entire teaching staff for additional feedback.

At the end of the semester, the committee presented the final results of the surveys and made recommendations to the principal regarding findings of the committee relative to the current evaluation system. In the summer of that year, the principal of Central High accepted the position of superintendent of schools in another state. A new principal was hired who, in turn, hired me as his assistant principal. He showed an interest in using the results of the study.

In the fall of 1987, at the beginning of the new Central High principal's first year, the diocesan central office staff, the principals of the Catholic grade schools in the Diocese, and the principals of the two high schools in the diocese began meeting to develop a standard system of

evaluation for the schools in the diocese. Input from administrators, elementary teachers, and high school teachers was gathered. Ultimately, the diocesan school board adopted a standard evaluation policy and accompanying procedures. Student ratings of teacher instruction was included as a part of information to be gathered in the evaluation process that was developed and adopted for use in the diocese by this group.

The first instrument used by Central High faculty for student ratings of instruction was a 19-item instrument developed by Manatt at Iowa State University and used in the state mandated evaluator training series. The teachers were given enough instruments and answer sheets for two sections of students. In spring 1988 the teachers chose any two sections of their classes and administered the instrument and returned the results to the principal's office. The results of the ratings for each teacher were tallied and a summary of the tabulated and written comments were returned to the teacher. This feedback was part of the discussion during the summative evaluation with each teacher at the end of the year.

At the end of the second year of this procedure, the principal received a request from a number of the teachers to make some adjustments to the instrument being used for the ratings by the students. They felt the instrument was

too vague and did not cover some areas at all. The principal asked the entire faculty for suggestions regarding additions or changes to the instrument. After all of the input was collected, he created a 40-item instrument that the faculty agreed was more comprehensive than the first. That instrument has been used since the spring of 1990. It is the results of those three years of student ratings that are the data of this study. (See Appendix A.)

The same procedure of collecting student ratings data and sharing it with teachers has continued each year since the spring of 1987. In the summer of 1992, as I was preparing to start a project for a graduate level class, the principal suggested that I use the results of the student evaluations to study for my project. He and I had an interest in what insights, if any, the results may have for us and the teaching staff. He had recently finished his Ph.D. and his dissertation was related to use of evaluation in school districts. He said he would be interested in finding out more about the area of student ratings. I had been involved in researching literature on teacher evaluation and found that student rating of instruction was often mentioned as a part of the information gathered for evaluation of instruction. The idea of finding out more about the topic was intriguing and could be useful for the



school, so I decided to create a study of the results of our student ratings of instruction.

After discussion with the principal, we decided that what might be valuable to the school is a study of the results of student ratings of instruction related to specific components of effective instruction. After the adoption of the evaluation system mentioned earlier, the teachers received considerable inservice training in the area of effective instruction during the 1987, 1988, and 1989 school years. A common language had been developed and was being used in the pre and post observation conferences. The language was included in the ratings instrument as well. The principal informed the faculty of our intent to look into the results of the ratings over the past three years and asked them for input then and along the way. The sentiment among the faculty was positive for such a study. Several have expressed an interest in the results.

In the literature of effective teaching, I have found a number of common components of effective instruction that were listed by several authors.

Rosenshine and Furst in "The Use of Direct Observation to Study Teaching" (1973) listed the following as effective elements:

1. Review (effective teachers begin a lesson with a 5 to 8 minute review)

2. Presentation of new material by
  - stating lesson goals
  - focusing on thought
  - teaching in small steps, and check for understanding before going on
  - give step by step directions
  - model the behaviors
  - organize the materials, so one step is mastered before going on
  - avoid distractions
3. Guided practice (teacher supervises students' initial attempt at a skill)
4. Provide feedback and correctives
5. Conduct independent practice (students working at a skill on their own)
6. Use weekly or monthly review

Berliner's "The Half-Full Glass: A Review of Research on Teaching" (1984) lists the effective components as:

1. Pre-instructional factors
  - A. content decisions.
  - B. time allocation decisions
  - C. pacing decisions
  - D. grouping decisions
  - E. decisions about activity structures

2. During instruction factors
  - A. engaged time
  - B. time management
  - C. monitoring success rate
  - D. academic learning time
  - E. monitoring
  - F. structuring
  - G. questioning
  - H. wait time
  - I. summary
3. Climate factors
  - A. expectations for achievement
  - B. environment for work
  - C. management of deviance
  - D. cooperative learning environments
4. Post instructional factors
  - A. tests
  - B. grades
  - C. feedback
  - D. evaluation

Madeline Hunter's "Decision Maker" model (1979) lists the following as effective components:

1. Anticipatory set (review prior learning and tie to the present)
2. Statement of objectives

3. Input (teacher giving information)
4. Modeling
5. Checking for understanding
6. Guided practice
7. Independent practice

Several of the components common to each author are included in the student ratings instrument used at Central High. I went through the effective components mentioned by the authors and chose five components that I knew, from past inservice with Central High teachers, would be familiar to staff. I then went to the ratings instrument and found two items from the instrument for each component of effective instruction I had chosen. Each item either addressed the component directly or was closely related to the component. In the end, 10 items were chosen from the instrument, 2 for each component of effective instruction. (Appendix B.)

The 10 items and the component of effective instruction that each is based are:

<u>Component</u>	<u>Item from Instrument</u>
Guided Practice	- I get enough practice with a skill or area before the class moves on to a new skill or area
	- If I don't understand, the teacher finds a way to explain so that I do understand

- Independent Practice - The teacher tests often enough to let me know how well I understand the material
- The teacher uses different ways (i.e., homework, questions, projects, written reports) to measure how much I know
- Teaching to Objective - The teacher discusses objectives and expectations of the course
- At the beginning of the class, the teacher tells me what activities we will be doing in the class
- Check for Understanding- The teacher encourages questions if the material/lesson is not understood
- The teacher gets responses from the students to determine if they are learning during the lesson
- Modeling
- The teacher's teaching methods encourage me to maintain interest during the class
  - The teacher explains difficult ideas in more than one way

In the instrument, the scale for each of these items had five choices available for the students to choose from for a response. For each of the 40 items, they were instructed to choose the one that best describes their teacher, what their teacher does, or the class. The responses the students had to choose from were: A = Always, B = Often, C = Sometimes, D = Rarely, E = Never

The next decision was related to selection of a sample of teachers whose ratings would be used for the study. Of the 62 teachers at Central High, 35 who were employed throughout the three-year period of 1989-90, 1990-91, 1991-92 had all administered the same rating instrument to their students. Included in the sample are a variety of teachers from several who started their first year of teaching in 1989 through teachers with over 35 years of teaching.

I tallied the student ratings from the 40-item instrument for each of the 35 teachers on a spreadsheet. A table on the spreadsheet for each teacher contained: the number of total student responses for each item, the number of student responses for the five choices available for each item, and the percentage of the total responses for each of the five choices offered. I calculated an overall average of responses for each of the five choices as well as the overall average percentage for each of the five choices. This study includes the results of ratings for the years

1989-90, 1990-91, and 1991-92. I used only the results for these three years because the same instrument was used over the three-year period.

A sample of the spreadsheet is:

Ques	Tot Rsp	A	A%	B	B%	C	C%	D	D%	E	E%
1	50	25	50%	10	20%	10	20%	5	10%	0	0
.											
.											
.											
40	50	10	20%	5	10%	25	50%	5	10%	5	10%
Ave	Tot	N	%	N	%	N	%	N	%	N	%

After all the teachers' ratings were in spreadsheet form, I separated the 10 items that corresponded to the five components of effective instruction from the rest of the items for each teacher and then studied the results of each teacher in the entire sample. Teachers whose ratings showed an increase over the three-year period were separated from the group. The five teachers whose ratings showed the greatest increase over the three year period were chosen for further study.

I interviewed these five teachers. The purpose of this exploratory interview was to determine, from the teacher's perspective, why the ratings increased in the areas listed. The following questions were developed for use in the interview. A brief rationale accompanies each question.

### Familiarity/Time

1. Are you familiar with the instrument used by students to rate your teaching near the end of each of the past three years?

This was asked to make sure that the teacher remembered the instrument and the time of year it was administered. Some teachers may use their own instrument to collect different kinds of information throughout the year and this way any confusion would be eliminated as well as focusing on the topic at hand. It was intended to give the teacher a chance to begin a general discussion of the topic of student ratings of instruction.

### Context of Class

2. You were asked to give the instrument to two of your sections. Do you have any thoughts on the number of sections teachers were required to administer the instrument to? Would you prefer more or less? Why?

In the literature, there was some discussion of the number of students in each section for student ratings. I was interested in the teacher's thoughts regarding how many students were in their sections and how and if they felt that affected the ratings.

3. Why did you choose the two classes you did to administer the instrument to?



I thought this question would help me understand the teacher's reason for choosing the classes they did since they were free to choose any of the sections they had to administer the instrument. A comparison of reasons and reasoning across the five teachers may prove to be interesting.

4. Do you remember anything significant about any of the classes given the instrument? Explain.

This was designed to further probe the classes chosen and why they were chosen to get a context of the classes administered the instrument.

#### Reaction and Interpretation of Results

5. Do you remember anything significant about the results you received from the instrument? Explain.

I wanted to give the teacher a chance to respond to anything they saw that was noteworthy in the students' responses. I also wanted to see what their overall reaction was to the ratings.

6. Ten of the items which are based on effective teaching components were chosen from the instrument. Students rated you generally higher in most of these areas in each of the three years. Can you talk about why you believe this happened in the areas of guided practice, independent practice,

teaching to the objective, checking for understanding?

This question was to focus the teacher on the 10 questions and five areas to study closely. I wanted to find out if there was anything specific the teacher did in any of the areas or in general over the three-year period to respond to the ratings he or she received and therefore affect the ratings.

7. In the area of (teaching to objective, modeling, independent practice, guided practice, checking for understanding) (choose one highest) the students rated you considerably higher each year. Why do you think this happened? Did you make adjustments or changes to make this happen?

If the teacher did not recognize the components of effective instruction, I wanted to point them out and probe the reasons for change they felt existed.

#### Response to Ratings

8. Did you do anything differently in your classes in subsequent years because of results of the ratings of any of the items on the instrument?
9. If you did something different, did you notice a change in your student's ratings in that area the following year?

Both of the questions above were meant to find out if the teacher used the ratings for feedback and to what extent they used them. I also wanted to see what his or her beliefs were regarding the effectiveness of their efforts.

#### Reactions to Ratings

10. Did/do you feel the results of the ratings were helpful or informative? Explain.

I wanted a general impression from the teacher of the usefulness of the student ratings.

#### Perceptions About Relationship between Ratings and Achievement

11. What function does student rating of teacher performance have relative to student achievement in your classes?

In the literature, this topic was discussed at length. I wanted to see if these teachers who had particularly significant increase in their student ratings had any thoughts about the relationship between their students' achievement and the ratings they received.

## Chapter 2

### REVIEW OF THE LITERATURE

This review cites selected literature that focuses on the involvement of students in the evaluation of teaching in the classroom. The review of selected literature relevant to this topic is organized in the following manner:

(a) general introduction with information related to the concerns regarding student ratings of instruction, (b) information related to resolving concerns with the process of using student ratings of instructors in the evaluation of teachers and their teaching, and (c) general findings and conclusions of the studies examined.

Many different aspects of evaluating teaching have been studied with mixed reviews from those that do the evaluating and those that are evaluated. An area that invokes considerable discussion from both sides is student evaluation of teacher performance. Aleamoni (1981) conducted an extensive search of studies related to the use of student ratings of instruction spanning the last 56 years. He found that student ratings can be useful in the evaluation process because:

1. Students are the main source of information about the educational environment

2. Students are the most logical evaluators of student satisfaction and effectiveness of course elements
3. The student-teacher communication involved raises the level of instruction
4. Use of student ratings by other students in course selection may increase the possibilities that excellence will be rewarded

Aleamoni used a considerable amount of research to substantiate his findings, yet student ratings receive mixed reactions from instructors for a variety of reasons. It is due to the inability of research to clearly resolve these concerns that student ratings of instruction remain a source of controversy.

When the concerns regarding student ratings of instruction are considered, there are eight that seem to be mentioned most often in the literature. The reasons that Aleamoni (1987) lists are used as the basis for the following discussion.

#### Discussion of Reasons

REASON 1. Students cannot make consistent judgments concerning the instructor and instruction because of their immaturity, lack of experience, and capriciousness.

The assumption is that students are too young and immature to make judgements about their instructors or the instruction they are receiving. Aleamoni found in his search of literature regarding the validity of student rankings, that student rankings tend to be pretty stable. Cohen (1981), in a study of 41 independent validity studies on 68 separate multisection courses relating student ratings to student achievement found that it is safe to say that student ratings of instructions are a valid index of instructional effectiveness.

A similar study by Sullivan and Skanes (1974) examined the correlation between student ratings and the amount students learned from instruction based on the score of a final achievement test at the end of a course. They were interested in whether positive correlations between grades and student evaluations existed on a more objective measure than a grade. They wanted to know if students could identify from which instructors they learned the most, and what qualities differentiated the successful and highly evaluated instructor from the less successful and or highly evaluated instructor. The information received would indicate not only the validity of the ratings, but also what kinds of characteristics the successful instructor exhibited.

Their findings showed that instructors who were academically successful and highly evaluated were both task oriented and interest arousing. Instructors who were unsuccessful but highly evaluated attempted to arouse interest without being task oriented. They also found that students' selection of subsequent courses in the same department were often affected by how well they achieved in the first course from that department. It would appear that if ratings are valid, students who are challenged and successful will continue with their pursuit of a discipline.

In another study by Marsh, Fliener, and Thomas (1975), a comparison of students ratings of instructors and student performance on a standardized final examination showed a positive correlation. It would seem that the consistent element across the three studies is that student ratings are a valid indicator of effective instruction. Marsh et al. found instructors attain improvements in instruction when they receive feedback from evaluations.

Probably the best evidence of students' ability to judge instruction is that they are the only ones who are in the class each and every day. From their experiences, they become "professional teacher watchers" and can make sound judgments from their experiences (Miller, 1988).

REASON 2. Only colleagues with excellent publication records and experience are qualified to evaluate their peer's instruction.

In 1973, Aleamoni and Yimer conducted a study by examining the publication records of the entire faculty of the University of Illinois, asking the faculty to rate the instructional effectiveness of the other instructors in their departments, and collecting ratings information from the students of instructors and courses. They wanted to see if they could determine the correlation between student ratings of successful faculty and faculty ratings of successful faculty. They found a high correlation, 0.70, casting doubt on the assumption that students cannot identify effective instructors.

In studies completed by Doyle and Critchton (1978) and Centra (1975, 1978), evidence was found that the reliability of student ratings is higher than that of colleagues.

REASON 3. Student ratings are nothing more than popularity contests with the warm, friendly, humorous, easy-going instructor emerging as the winner every time.

This reason contains issues that are probably the most difficult to decipher. The characteristics mentioned are also often mentioned as successful characteristics of an effective teacher (Brophy & Good, 1986). It may be



difficult, therefore, to separate the positive qualities mentioned from the other criteria of effective instruction being measured. Freeman (1988) considered the successful relationship between counselor and client to be predicated on the client's perception of the attractiveness, expertness, and trustworthiness of the counselor. His study assumed that the characteristics that are associated with perceived effective counseling and teaching are similar. He likened the teaching process to the counseling process in that both involve human relationships. It would seem that one aspect of receiving high ratings would be if the relationship between the student and instructor is a positive one. Common sense would indicate that most students would rate a person who was warm, friendly, and humorous high. Freeman's 1988 study indicated that perceptions of students on the Counselor Rating Form-Short version are positively related to judgments of teacher effectiveness but further research is needed to find out what influences this correlation.

This argument implies that students are not able to make judgments about specific aspects of a course or make critical judgments among a variety of characteristics. Costin, Greenough, & Meneges (1971) refer to this as the "halo affect" in their study. The "halo effect" implies that students give good ratings to instructors who are

entertaining and who deliver courses devoid of content. The study found that this just does not happen. For the most part, students make sound, meaningful judgments about courses.

A similar phenomenon is referred to as the Dr. Fox effect. This effect is related to the influence of the instructor's personality on student ratings. Supposedly, the more expressive and entertaining the instructor, the higher the ratings no matter what the level of meaningful content in the lesson. Marsh and Ware (1982) found that students may rate an instructor high for enthusiasm, but do not confound this with areas of course content or other effective elements of instruction. A key part of the study found that the ratings of enthusiasm of an instructor are affected by the motivation of the student. If the student is given a reason to learn the material, how enthusiastic the instructor is has less of an effect. Enthusiasm can be higher if the instructor is entertaining but has little solid content to share. The students respond more to the entertainment because that is what they have left to respond to if no real content is being presented.

REASON 4. Students are not able to make accurate judgments concerning either instruction or instructor until they have been away from the course, and possibly away from the institution, for several years.

The assumption here closely resembles the argument that students are too young and immature to accurately judge effective instruction. Actually, this argument has to do with students use of what they have learned. A number of studies of alumni ratings (Aleamoni, 1974; Centra, 1974; Drucker & Remmers, 1951; Marsh, 1984; McKeachie, Lin, & Mendelson, 1978) have found that when ratings of instructors change after graduation, they do not change dramatically enough to be significant.

REASON 5. Student rating forms are both unreliable and invalid.

Teachers use a variety of rating forms for a variety of reasons. Research has yet to produce a common rating form that could be used across the disciplines or departments. Because of this, rating forms have been questioned and, in many cases, probably with some legitimacy. Newstead and Arnold (1989) attempted to determine what kind of responses used on a rating seemed to be the best. They had students rate their lecturers using three different types of responses: the verbal label, the end anchored, and the numerical scale. If accuracy of measurement was desired, the authors found that the numerical scale was the most accurate. The reason was the numerical scale was more precise and the students were able to delineate more clearly what they wanted to indicate for a rating.

Kemp and Kumer (1990) make several recommendations regarding teacher satisfaction with ratings instruments. They encourage use of a well constructed, proven instrument, not one that was just made up by the staff. An appeal process should be built in.

A study by Cohen (1981) of the relationship between student ratings of instruction and student achievement found that ratings of instruction are a valid indicator of how effective instruction was. He felt that students do a good job of being able to determine which teachers are effective. His findings are supported by Lamberth and Kostaschi (1981) whose study used student evaluations from two groups of students in the same class: one group would be receiving a grade from the instructor, and one would not be receiving a grade. They felt this would eliminate the pressure to give high ratings because of the influence of grades. They found that persons knowledgeable about the course content, who attended class regularly, read the text, and were exposed to the teaching of the instructor without the threat of a grade, evaluated instructors in similar fashion as do graded students in the class.

REASON 6. Extraneous variables or conditions such as size of class, gender of student or teacher, time of day course is offered, if course is a requirement or an elective, whether the student is in a major or

non-major course, level of the course, or rank of the teacher affect student ratings.

There are mixed findings in each of these areas but in most, there is not enough evidence to indicate that any of these has a profound effect on student ratings. Studies concerning class size do show that smaller classes usually result in higher ratings but it is because more learning usually takes place in smaller classes (Centra, 1979). What is worth noting when collecting student ratings is that ratings from classes of considerably different sizes should not be compared. Required classes generally get lower ratings. The reason could be required classes generally tend to be larger and therefore students do not feel that they have received the attention or rapport received in a smaller class. In fact, in a study by Toby (1988), the results showed that the smaller the class, the greater the learning. Toby felt that under the circumstances, for the instructor of a large class, it would be "remarkable for a teacher to get high ratings, yet some do!"

Kemp and Kumer (1990) found that a class could be too small for reliable ratings and put 10 as the smallest class size to use and still get reliable ratings. This study stated a valid point in that ratings are just indicators, they are not precise measurements and teachers and administrators should interpret them accordingly.

REASON 7. Grades or marks that the students either expect to receive or actually receive are highly related to their ratings of both course and instructor.

Aleamoni (1987) compared the research from a number of studies and found that there was almost no correlation between grades and ratings. He stated that grades are notoriously unreliable and do not necessarily reflect what the students have actually learned.

REASON 8. There is no way student ratings or evaluations can possibly be used to improve instruction.

In order for student ratings to be useful in this category, they must assist in accurately diagnosing problems or prescribing a solution to problems. McKeachie (1979) studied student ratings of faculty to determine if student ratings measure teaching effectiveness. In the study, he drew distinctions among the different reasons student ratings are gathered, one of which was improvement of instruction. From his study, the key element for improvement of instruction relative to ratings information was that the ratings must be shared with the instructor and discussed so the ratings become positive feedback. Pambookian's (1974) research found that instructors in the middle range of ratings tended to benefit from feedback (student ratings) while the top and bottom teachers did not,

suggesting that teachers receiving low ratings may become discouraged. Overall and Marsh (1977) found that students of instructors receiving feedback from student ratings gave their instructors higher ratings at the end of the year and scored higher on an achievement test.

If the evidence indicates that student ratings are effective for improvement of instruction, why then does some teachers' instruction fail to improve after feedback?

McKeachie suggested the following:

1. The ratings may not provide new information
2. Low ratings and critical comments may create anxiety, discouragement, and lack of enthusiasm for teaching-lowering rather than improving motivation for teaching
3. Even when faculty members want to improve, they may not know what to do

Teaching is a human process. In a human process, it would seem natural to give and receive feedback. Riley, Bryce, and Lifshitz (1987) suggest that teaching is such a human endeavor that getting some feedback, even though less than perfect, is valuable to see how what one is doing is being received. A sort of "need to know" for the consumer. Stevens (1987) goes on to suggest that student ratings (feedback) can help improve instruction if accompanied by support of the cognitive needs of the instructor. The

cognitive needs are motivation, attitude, and knowledge. Evaluation of teaching is complex and simply relaying feedback is not enough. Assistance for the instructor on how to conceptualize, interpret, and apply feedback is vital (Stevens, 1987).

### General Findings and Conclusions

It appears from the research that the issues regarding the use and interpretation of student ratings are far from resolved. Considerable evidence does seem to exist, though, to support the continued use of student ratings as a part of the evaluation of instruction. If ratings are used, the purpose or intent of their use should be clearly defined and clearly explained in advance. Realizing that ratings are only indicators, not precise measurements, will help keep their importance relevant.

Overall, the use of ratings seems to be a sound practice. If included into an evaluation system, greater success and satisfaction will be achieved if: ratings are part of a well defined system, are interpreted by objective and fair evaluators who are knowledgeable of the position evaluated, an appeal procedure is included, and teachers are positively encouraged to take suggestions from the feedback and try to improve their instruction.



One point of interest is almost all the literature researched focused on college students' ratings of professors. Very little research was found regarding high school aged students rating their teachers. My study focused on high school students' ratings of their teachers. The results could be compared to other studies to determine if there are differences between the age groups and if so, if they are significant.

## Chapter 3

## PROCEDURES

I chose 5 teachers from the 35 teachers in the sample to interview. Based upon criteria previously mentioned (increase in percentage of positive student response over the three-year period in 10 selected areas on the rating instrument). These 5 teachers included one man and four women: Dave Meredith, a 34-year-old vocal music instructor with 12 years teaching experience who is currently working on a masters degree; Donna Lynn, a 44-year-old theology teacher with 10 years teaching experience; Joan Donlin, a 28-year-old art teacher with 7 years teaching experience; Jo Reese, a 37-year-old English teacher with 15 years experience; and Karen Holt, a 33-year-old French teacher with 12 years teaching experience.

The interviews took place within a two-week period. I conducted each interview either in the teacher's classroom or office and recorded each on tape. I used the same set of questions for each interview. The interview questions are included in Appendix C. The teachers knew that I was working on a study of student ratings, including the current system in use at Central High, and I asked them to be open and candid during the interview. I made a written

transcript following the interview and gave a copy to the teacher.

The interview questions were developed around consistent themes or concerns from the literature that teachers mentioned most often regarding student rating of instruction and instructors. I found it interesting that the teachers I interviewed were familiar with these concerns. In the literature I found few studies or references to studies that included feedback from high school teachers about high school students' ratings. Because of this, I was not sure what kind of responses I would get from high school teachers. It seemed that each question struck a familiar concern that the teacher had an opinion about. Most were able to express their ideas at some length and felt they had experiences in the classroom related to the question. Apparently, from my interviews at least, commonalities exist between post-secondary instructors and high school teachers regarding student ratings.

Personal educational philosophy was often laced throughout the answers in what I thought was a sincere effort on the teachers part to make their responses uniquely clear from their perspective. Most shared personal experiences with students or classes and explained how these specific incidents were taken as indicators of valuable

insights into the level of their relationships with the class or students, or of the degree of effectiveness they have in teaching the subject or relating to the students. Most saw the positive ratings as reinforcement for efforts they were making and negative ratings as indications of what needs work.

Since each of these teachers had ratings that seemed to improve over time, one would assume that the teachers were being effective in the classroom and with the students. The level of sincerity toward their subject matter and the students was consistently high across the group. The level of real effectiveness in teaching the students the material and recording sustained or real growth in skills in the particular areas was not as easy to determine. Because each of these teachers is generally perceived by the students in school as a good teacher, I wonder if the "Halo Effect" may be present for each of the members of the group? I think some of that exists, but more variables are involved.

For instance, there is considerable difference in the planning, approach, and level of expectation for each of the teachers in the group. All of the teachers do a good job in the classroom and have good knowledge of the subject matter. One of them could be labeled a master teacher on the following basis: the students in the class consistently do outstanding work and show remarkable determination and

growth. Each year, a large number of students reach all-state status, a few are accepted into somewhat prestigious music programs, and the overall performance of the classes in general are evident at public performances. This teacher is very disciplined and demanding and yet showed considerable increase in positive responses from the students over the years. It is quite possible that the students truly appreciate the skills and knowledge learned during the class and understand that the hard work that is expected will pay off. Their identification on the survey of the amount and level of effective teaching practices used seems to reinforce the fact that learning takes place.

Other teachers in the group also have expectations for their students. They seem to focus more on the student-teacher or student-school relationship. These are excellent teachers also. Each of them focused more on the student and their relationship to the class or teacher than on the skills acquired. Their ratings from the students over the three years were also high and it appeared that the students perceived them to be quite effective. One of them in particular is very popular with the students. During registration, many students request her as a teacher, and some students are often turned away because of lack of space. Her expectations are of a different nature than the teacher mentioned above and the students respond very

positively to her. They enjoy her class and seem to grow from the experience. This teacher would take very seriously any student who was not happy with the class and would work hard to change whatever it was to make the relationship between them better.

Because of what I knew about each of the teachers, I thought that a number of very different themes would develop during the interviews regarding student ratings. I did find some differing opinions, but was surprised by the amount of similarity that existed within the answers the teachers gave. Following is a list of themes taken from the answers of the teachers in the interviews:

Dave:

1. I don't fear evaluations but others do
2. The feedback proved useful to me
3. Some of the questions just don't pertain to me
4. We should work by departments to generate items
5. Not sure the kids can make good judgments
6. More accurate assessments of courses appear later,  
a few to maybe 10 years down the road
7. Specific leanings

Joan:

1. Kids negative because have to do some many at same  
time
2. If done at mid-year, teacher could change

3. Appreciate the written comments the most
4. Some questions don't pertain
5. Kids are truthful, good judges
6. The time of day makes a difference
7. Kids here are more sophisticated about what they need or want
8. Give kids a chance to include what they want from the department in the evaluations

Donna:

1. Evaluate each semester so can incorporate into the next
2. Current system makes us accountable
3. Some questions don't pertain, let department make
4. Needs to be done more than one time a year
5. Kids are accurate judges most of the time and have their finger on what is good teaching
6. Like written comments the most
7. Good to give to classes and compare
8. Administration should sit down with teacher and talk about ratings the teacher gets
9. Give more open-ended questions

Karen:

1. I do my own evaluation each semester
2. Need more questions of kids about how class went

3. Spread evaluations out, like social studies one month, etc.
4. Written comments are most useful, give more room for them
5. Kids are good judges, just need to do in neutral situation
6. Do a pre and post evaluation of when student enters class and then after they are in college
7. Get counselor feedback also, kids talk to them
8. Go over results with teacher in the post observation
9. Have the student council make up some questions

Jo:

1. Do my own, use to make adjustments in class
2. Don't have students do all teachers at one time
3. Do more than one time a year
4. Students are pretty good judges
5. Like it when students with low grades give high ratings, means I connected
6. Like the written comments, remember them more than others
7. Find evaluation very uncomfortable but helpful
8. Include student evaluation with administration's evaluation



9. It is important for students to know they have a say through the ratings, but do not confuse style with substance

## Chapter 4

## ANALYSIS AND INTERPRETATION

After studying the results of the interviews, I believe that one of the common themes mentioned throughout was the teachers' interest in the make-up of the items in the rating instrument. Each teacher felt that a portion of the items did not pertain to them or their area and therefore, that portion was not valid or useful in obtaining useful information. When asked what suggestions they might offer to improve the instrument, several teachers said that departments should be given the task of working together to come up with a series of questions that would be more specific to their area.

Dave felt there would be a much different survey and the results would be more valuable because teachers would be able to find specific areas the students are not excelling in or places the students are not grasping the information, recognize these as strengths or weaknesses, and then work on them. Dave does an evaluation of his own each year that includes areas suggested above. He combines the results with those received from the yearly survey completed by the students at the direction of the administration. The items from his evaluation that match the items of the school-wide survey seemed to have a high correlation he believed. As

mentioned, Dave is working on his masters in vocal music. He is an excellent teacher, perhaps one of the best at Central High. He is very disciplined in his approach to planning his program and the students in his choirs are always well prepared and do an excellent job. The students seem to enjoy the classes, even though they are sometimes demanding, and from most indications, seem to feel a real sense of accomplishment and pride from being associated with the program. Dave's answers were indicative of the type of person he is and the type of program he runs. He looked at the items as something to use as feedback to concentrate on improving the program. His concern was with the skills the students were learning. He wondered if were they being taught effectively, and if he could learn something from the students feedback to help him get them to do better. Dave has excellent rapport with his students but he was not as interested in the survey giving information about the relationship between him and the students as he was about ways to improve the program. This particular approach was not the same for all the teachers interviewed.

Joan stated that some of the questions do not pertain to her as an art teacher. She is currently enrolled in a collaborative education class and the topic has been assessing what the students' needs are in the class. Joan felt that questions about the broader picture in school

would make the instrument better. She believes that questions about what could be better addressed by the Art department or school in general should be included. She and other members of the art department often share ideas and make comments about what is going on in each other's classes. Joan feels this kind of casual exchange also occurs with students and helps her evaluate the class and her teaching. A sense of "how is it going" seemed to be, either in a formal or informal way, an important ingredient necessary for her results. She, too, was interested in the skills the students were learning but she was also quite interested in the opinions the students had about the class or department in general.

Like Dave, Joan has excellent rapport with her students and is confident about her relationship with them. The emphasis she seemed to have relative to the items was in the area of how the school or department was connected with the student. This seemed to put the emphasis on the content and procedures used, then on how the teacher could manipulate those to better suit the student's needs.

Donna states that the problem with questions that do not pertain to the class is that students just guess or make up answers to fill in the blanks. Karen also felt that some of the questions only gave very superficial information with the students answering almost without thinking, just filling

in bubbles or making designs on the response sheet. She felt, as Donna did, that students should be questioned more about things that are meaningful in the relationship between the teacher and the class. This information would be more helpful to the teacher than what is received now. I believe Donna and Karen feel strongly that they have good rapport with their students and that their students like them as teachers and people. It may be this characteristic that motivates their desire to increase the amount of input about what students think about how things are going in the class between the teacher and the student. Each seems to have a strong need to have a positive connection with the students. In their classes, students feel a strong personal connection with them. In some cases, Karen has the same students for two or three years because she teaches a sequence of the language. She often refers to how well she and the students know each other and the expectations they have for each other. At times, it is as much like a relationship between friends.

This same type of relationship is especially true for Donna. She teaches a class that naturally requires a closeness between teacher and student because of the amount of sharing that takes place about topics which are frequently of a personal nature. Donna tends to be extremely sympathetic with her students and it is important

to her that they feel good about themselves, the subject, and her. Compared to Dave, she seemed less concerned about the strategies of effective teaching and more concerned about the affective domain in the classroom. In the interview, she stated that the students may not be able to verbalize what is good teaching, but they know it when they see it. She also states that the students tend to be very positive with her and it appears she works hard to make sure that kind of relationship exists in the class all the time. She wants the students to like her and her teaching style. Her techniques support that effort. It appears that she recognizes effective teaching components but believes that building a nurturing relationship between student and teacher is equally necessary.

Jo had little to say about the questions that do not pertain. She did mention that the written comments the students made were the most meaningful to her, a comment that was also made by each of the other teachers. The teachers seemed to take great interest in what the students had to say. They all reported this was the first thing they looked at or read when they received their results. Several recommended more space be provided for this type of answer when asked for suggestions to improve the current instrument.

The written responses that the teachers received were not recorded for this study, but from their comments I judge that most of the written comments received were positive in nature. Teachers believed that the students were taking time to offer their own thoughts and this somehow carried special meaning for them, something the stated items on the survey could just not convey. Karen and Jo did mention receiving negative comments. They both seemed to think that the student may have had a grudge or that they were mad over a grade or something that happened in school. Therefore, this may not have been a valid answer relative to rating how the course went in general. Karen suggested students be placed in a "neutral" situation when evaluating a class or teacher. This may then eliminate or at least diminish the skewed effect of the response. I am not sure how one would create a neutral setting or condition. Having students with a negative answer may be a necessary part of the process.

The teachers did not dwell on the connection between the items on the instrument and the characteristics of the effective teacher. Few reported studying the responses in some detail to determine the specific responses of the students to efforts the teacher may have been making in specific areas as stated by the instrument. The teachers seemed to have more significant interest in personal reactions or comments about the class or procedures. It

seemed that a more thorough review of the responses to each item in the instrument as related to quality teaching could be recommended. During the interviews, several of the teachers suggested administration should meet with them and discuss the results when they are collected. The procedure has been to return the results to the teachers and let them study the responses and draw their own conclusions, more as an informational step in the evaluation process. If a more meaningful understanding of the results is sought, it appears a change in the procedure of returning and interpreting responses is needed.

Several studies in the literature focus on the instrument used in student ratings. Kemp and Kumer (1990) spend considerable time discussing the need for a well constructed, proven instrument, not just one made up by the teachers or administration. I thought of this during the interviews when teachers suggested who should be involved in the creation of items for the ratings. They may have seen the items more as a means of evaluating something specific to their style or program instead of a general rating of the their use of effective teaching practices. Creating an instrument that was specific to a discipline could be construed as putting the traditional evaluator of teaching at a disadvantage because the evaluator may not be trained in the rudiments of the program or discipline, and therefore



not be capable of evaluating, rating, or interpreting how the class went or the job the teacher did. This would not be the case if both areas, subject matter, and effective teaching components, would be covered in the ratings. It may be beneficial for the department to agree on certain areas to receive feedback, the teacher to have a section for their concerns, and then the general teaching to be covered.

Several teachers indirectly mentioned the question of validity when they stated that some of the questions just did not pertain to them or their area. I took this to mean the teachers did not feel the questions contained the criteria of the course and therefore were seeking information that either may not or should not exist. Since there is no single criterion of effective teaching, their concerns are probably valid. As mentioned above, this may be something that needs to be addressed if it is desirable to have the teachers see this instrument as useful and valid. Considerable time has been spent in the past few years at Central High on what are effective teaching characteristics. Observations of teachers have continued to draw attention to the use of these characteristics, but inservice may have to be included to focus once again on their importance. This may be the time to include the make up of student ratings and how they reflect the use of the characteristics if that is what we want them to be used for.

What seems important to remember is to be realistic with what we expect students' ratings to do. When teachers look to the written comments received from the students or question how the certain items pertain to them, what they may be saying is that there is some information being shared that no one in the administration of the process knows about or was able to anticipate, but can be found useful. What may need to happen is for all parties to decide what components of the teaching process each wants feedback on and then decide on the best way to ask for each. Since most of the information that is being sought does not have exact description, the degree of precision sought from the responses should be interpreted with care. This may have something to do with the teachers' interest in the students free responses to items or in general in the ratings.

The type of rating scale used in the current instrument may have some influence on the interpretation of the items by the teachers. Newstead and Arnold (1989) indicate that there may be a difference in the accuracy of results of certain types of format responses. The Central High instrument uses the verbal scale of, A = Always, B = Often, C = Sometimes, D = Rarely, and E = Never. Along with verbal scales, responses to ratings from two other types of scales, end anchored, in which only the end points are given labels, and numerical scales, in which the points are numbered but

not given verbal descriptions were studied. The point of the research was to compare the results obtained. The comparisons were to see if the scales differed in any of the following ways:

Overall Means - does one scale produce higher ratings

Reliability - does one produce similar results as  
repeated

Validity - is it accurate (measures what says it does)

Interval Scaling - are points psychologically  
equidistant

Variability - are responses spread along the scale

The study found that there was little or no difference between verbal and end anchored scales. The end anchored scale tended to attract responses to either end of the scale, allowing less variability. The numerical scale produced responses different from the other scales. This scale produces what seems to be a greater degree of accuracy because respondents are able to choose a more precise rating.

This may explain why our teachers may not have confidence in the ratings they receive. If the results are vague or left to an interpretation that is not precise enough to make sense, teachers will be frustrated. The open-ended or written comments teachers receive are probably looked upon as more precise. We may want to explore

changing our scale to one that is numerical with precise percentages at points along the scale. This will give more precise information and because they are numerical, the response values can be used for comparison or study. We could add more space for students to give written answers as well to help explain the percentages marked.

Frey (1976) observes that a decision should be made in advance whether the students are being asked to be reporters or judges. He notes that results are more valid when students are asked to make observations rather than state general feelings of like or dislike for the course. The latter puts students in the role of evaluator and the validity may then be suspect. From the interviews, it appears that the student as judge may have been the most desired or expected from the teachers, or at least they were very interested in what the students' judgments were. This may be an area address so that it is clear what role our students are asked to play. One type of information may be received while being interpreted as another.

With this in mind, it seems important to be aware of the intentions that both teacher and student have regarding the practice of rating instruction and the use of results. Several of the teachers interviewed commented on the timing of the student ratings. The fear was that too many teachers are asking the same students at the same time of year to

rate their class or instructor. With the large number of ratings that students may have had to do at the same time, teachers felt that students did not take them seriously and therefore the results may not be not valid. We need to insure the students that these are meaningful exercises and that the results will be used by the teacher. If their perception is that these are going to be used, it seems they would take care in filling them out. If the procedure we currently use does not reinforce this, we should look at how to improve it.

A negative perception by the students can lead to a negative perception of teachers as well. Teachers may not put faith in and may actually feel threatened by students' ratings if they doubt the validity of the results. Teachers may not put much time or give much consideration to areas of their teaching the results say need attention if validity is in question.

Sullivan and Skanes (1974) addressed the issue of validity of student evaluations and the characteristics of successful instructors. They found the academically successful teachers tended to arouse interest and were "task oriented." Unsuccessful but highly evaluated instructors tended to arouse interest without being task oriented. Having successful students and receiving high evaluations did not seem to be automatically related. This study

reminded me of Dave and his classes. He is a task oriented teacher, receives high evaluations, and his students are successful. It would be my observation that his students take the opportunity to evaluate his teaching and the class seriously and he takes the results seriously as well.

The issue of student and teacher perception of the process may be the most important in attempting to insure validity because so many concerns enter into what influences these perceptions: are the items meaningful, are the reasons for doing the rating clear and communicated to all, is the connection between the items and what happens in the classroom understood, are the results valued, shared, and used appropriately, will the results be taken seriously, are the students aware of their role in the rating process, and are all the stakeholders involved in the creation and evaluation of the process?

As I stated earlier, it was my intent to see if the teachers in the study felt there was any relationship between their high ratings in the areas chosen and the achievement of their students. With several teachers, they vaguely remembered receiving feedback that resulted in a change of their procedures in the class. It was difficult for them to think of the particular things that students marked or that class results indicated. Easier for them to remember were the written statements or the general

impressions that the results contained. They were quite sensitive to the role of the students as judges and were genuinely interested in what the students had to say. I interpret this as indicating that these five teachers do feel that in general, student input is valuable to them and will in the end, influence what happens in their classrooms.

## Chapter 5

### SUMMARY AND CONCLUSIONS

In relating the findings of this study to the student rating process at Central High, I would make the following recommendations to improve the process:

Time of year - Teachers should be given the option of conducting the ratings either at the end of the first or second semester. New teachers should do ratings twice the first year or two, and veterans given the option of more than once if they want. This would allow teachers to get feedback and use it right away during the year. It would also diminish the problem of students being overloaded with a large number of ratings to do at a single time in the year. The richness and quality of responses would be preserved.

Instrument content - The instrument should be divided into parts. One part would rate teacher behaviors as described by effective teaching components. Another part would rate course content. This section would include questions directed at specific areas the department would like to ask students. Ratings with fixed responses could be shared with all members of the same department so that a department could evaluate certain areas collectively and make adjustments if necessary. Faculty ownership could be



enhanced with this procedure. These responses could also be used during the formal evaluation period with the administrator or with teachers who have mentors. A third part could be developed that asks the student to rate his or her learning behavior in the class. This would help reinforce their role and expectations in the class and let them see they are a part of the learning/teaching process.

Data gathering - A data base should be developed so that students could enter their responses directly into a computer so that compilation of results would be faster and be tracked over a period of years. A procedure of keeping students' answers anonymous yet being able to track a disgruntled student may be beneficial. If a teacher could get a printout of the results of his or her classes in a matter of a few days, the results may be seen as more useful and meaningful.

Selection of students - All students should be asked to do ratings each year for all of their classes. This would give a complete picture to the teacher.

Inform students - Administration should present the ratings process to freshmen in small groups or classes. This would help students understand from the beginning the make-up of the instrument, procedures used, expectations of students and teachers, and uses of the feedback. A

consistent explanation and value would be given to the process.

## REFERENCES

- Aleamoni, L. M. (1974). Typical faculty concerns about student evaluation of instruction. Paper presented at the Symposium on Methods of Improving University Teaching. Haifa, Israel: The Technion Institute of Technology.
- Aleamoni, L. M. (1981). Student ratings of instruction. In J. Millman, ed., Handbook of teacher evaluation (pp. 110-144). Beverly Hills, CA: Sage.
- Aleamoni, L. M. (1987). Typical faculty concerns about student evaluation of teaching. New Directions for Teaching and Learning, 31, 25-31.
- Aleamoni, L. M., & Yimer, M. (1973). An investigation of the relationship between colleague rating, student rating, research productivity, and academic rank in rating instructional effectiveness. Journal of Educational Psychology, 64, 274-277.
- Berliner, D. (1984). The half-full glass: A review of research on teaching. In J. Hosford, ed., Using what we know about teaching (p. 54). Alexandria, VA: Association of Supervision and Curriculum Development.
- Brophy, J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), Handbook of research on teaching (3rd ed.) (pp. 328-375). New York: Macmillan.
- Centra, J. A. (1974). The relationship between student and alumni ratings of teacher. Educational and Psychological Measurement, 34, 321-325.
- Centra, J. A. (1975). Colleagues as raters of classroom instruction. Journal of Higher Education, 46, 327-337.
- Centra, J. A. (1978). Using student assessments to improve performance and vitality. New Directions for Institutional Research, 20, 31-49.
- Centra, J. A. (1979). Determining faculty effectiveness. San Francisco: Jossey-Bass.

- Cohen, P. A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies. Review of Educational Research, 51, 281-309.
- Costin, F., Greenough, W. T., & Meneges, R. J. (1971). Student ratings of college teaching: Reliability, validity, and usefulness. Review of Educational Research, 41, 511-535.
- Doyle, K. O., & Crichton, L. I. (1978). Student, peer, and self evaluations of college instructors. Journal of Educational Psychology, 70, 815-826.
- Drucker, A. J., & Remmers, H. H. (1951). Do alumni and students differ in their attitudes toward instructors? Journal of Educational Psychology, 42, 129-143.
- Freeman, H. R. (1988). Perceptions of teacher characteristics and student judgements of teacher effectiveness. Teaching of Psychology, 15(3), 158-160.
- Frey, P. W. (1976). Validity of student instructional ratings. Journal of Higher Education, 47, 327-336.
- Hunter, M. (1979, October). Teaching is decision making. Educational Leadership, 37, 63.
- Kemp, W., & Kumer, G. S. (1990). Student evaluations: Are we using them correctly. Journal of Education for Business, 65(8), 106-111.
- Marsh, H. W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential bias, and utility. Journal of Educational Psychology, 76, 707-754.
- Marsh, H. W., Fliener, H., & Thomas, C. S. (1975). Validity and usefulness of student evaluations of instructional quality. Journal of Educational Psychology, 67, 833-839.
- Marsh, H. W., & Ware, J. E. (1982). Effects of expressiveness, content coverage, and incentive on multidimensional student rating scales: New interpretations of the Dr. Fox effect. Journal of Educational Psychology, 74, 126-134.
- McKeachie, W. J. (1979). Student ratings of faculty: A reprise. Academe, 65, 384-397.

- McKeachie, W. J., Lin, Y. G., & Mendelson, C. M. (1978). A small study assessing teacher effectiveness: Does learning last? Contemporary Educational Psychology, 3, 352-357.
- Miller, R. I. (1988). Evaluating faculty for promotion and tenure. San Francisco: Jossey-Bass.
- Newstead, S. E., & Arnold, J. (1989). The effect of response format on ratings of teaching. Educational Research and Psychological Measurement, 49(1), 33-43.
- Overall, J. U., & Marsh, H. W. (1977). The relationship between students' evaluation of faculty and instructional improvement. Paper presented at Third International Conference on Improving University Teaching, New Castle-upon-Tyne, England.
- Pambookian, H. S. (1974). Initial level of student evaluation of instructor as a source of influence on instructor change after feedback. Journal of Educational Psychology, 66, 52-56.
- Riley, J. W., Jr., Bryce, R. F., & Lifshitz, M. (1987). The theory and practice of student evaluations. In The student looks at his teacher (pp. 21-34). Washington, NY: Kennikat Press.
- Rosenshine, B., & Furst, N. (1973). The use of direct observation to study teaching. In R. Bravers (Ed), Second Handbook on Teaching (p. 130). Chicago: Rand-McNally.
- Stevens, J. J. (1987). Using student ratings to improve instruction. In New directions for teaching and learning (pp. 33-38). Berkeley, CA: Jossey-Bass.
- Sullivan, A. M., & Skanes, G. R. (1974). Validity of student evaluation of teaching and the characteristics of successful instructors. Journal of Educational Psychology, 66, 584-590.
- Toby, S. (1988). The relationship between class size and students' rating of faculty. Journal of Chemical Education, 65, 788-790.

APPENDIX A  
RATING INSTRUMENT USED BY STUDENTS

Student Evaluation of Instructor

Select the answer that best describes your teacher's teaching or what happens in class.

Scale: A = Always; B = Often; C = Sometimes; D = Rarely;  
E = Never

A. Diagnosing and Monitoring

1. The class moves at about the right pace for me
2. I get enough practice with a skill or area before the class moves on to a new skill or area
3. If I don't understand, the teacher finds a way to explain so that I do understand
4. The work is not too hard or too easy for me
5. The teacher tests often enough to let me know how well I understand the material
6. The teacher uses different ways (i.e., homework, questions, projects, written reports) to measure how much I know
7. The teacher tells me what to study in preparation for a test
8. The tests are based on what was covered in the course
9. Assignments and tests are returned promptly.
10. Returned assignments and tests have comments from the teacher which help me understand what I need to work on
11. The teacher reviews/discusses work to help me understand my strengths and weaknesses

12. Every day the teacher uses some way to determine what I need to know
13. The teacher has an accurate idea of my abilities and my knowledge

B. Decision-making

14. The teachers knows his/her subject area
15. The teacher adjusts the course to the students interests and needs
16. The teacher relates events that happen in life to the subject being studied
17. The teacher uses different grouping patterns (i.e., whole class, small group, pairs, etc.) depending upon the lesson
18. The teacher uses a variety of materials

C. Planning Instruction

19. The teacher discusses objectives and expectations of the course
20. The teacher discusses the objectives and standards for each lesson so that I know what is expected of me
21. At the beginning of the class, the teacher tells me what activities we will be doing during the class
22. The teacher explains how lessons are related to the previous lesson

D. Instruction: Teaching the Lesson

23. The teacher uses questions and responses from the students during a lesson
24. When students give wrong answers, the teacher helps them develop the right answers
25. The teacher encourages questions if the material/lesson is not understand
26. The teacher gets responses from the students to determine if they are learning during the lesson

- 27. The teacher uses more than one method of teaching (i.e., lecture, worksheets, student presentations, etc.) during a lesson
- 28. The teacher's teaching methods encourage me to maintain interest during class
- 29. The teacher explains difficult ideas in more than one way
- 30. The teacher acknowledges good work and responses from the students
- 31. The teacher leads discussion well

E. Managing the learning environment

- 32. The teacher helps all students to feel comfortable in the class
- 33. The teacher is fair with all students
- 34. The teacher avoids offending or putting down students
- 35. The teacher builds confidence and self-esteem in all students
- 36. There is fair and consistent discipline in the class

F. Home-school relations

- 37. When problems arise, the teacher works with me and if necessary with my parents to work out a satisfactory solution
- 38. The teacher uses parent and other adult volunteers in the class

G. Human relations

- 39. I feel comfortable going to my teacher with concerns or problems
- 40. The teacher shows interest in me and cares that I do well in class



# APPENDIX B CUMULATIVE RATINGS

## Teacher Donna Lynn

Yr	Q#	Tt	Rsp	A	A%	B	B%	C	C%	D	D%	E	E%
----	----	----	-----	---	----	---	----	---	----	---	----	---	----

### Guided Practice:

I get enough practice with a skill or area before the class moves on to a new skill or area.

89	2	22	15	68	5	23	2	9	0	0	0	0
90	2	20	18	90	1	5	1	5	0	0	0	0
91	2	35	33	94	2	6	0	0	0	0	0	0

If I don't understand, the teacher finds a way to explain so that I do understand.

89	3	22	13	59	2	9	6	27	0	0	0	0
90	3	20	16	80	4	20	0	0	0	0	0	0
91	3	35	30	86	5	14	0	0	0	0	0	0

### Independent Practice:

The teacher tests often enough to let me know how well I understand the material.

89	5	22	6	27	9	41	6	27	1	5	0	0
90	5	20	18	90	2	10	0	0	0	0	0	0
91	5	35	30	86	5	14	0	0	0	0	0	0

The teacher uses different ways (i.e., homework, questions, projects, written reports) to measure how much I know.

89	6	22	13	59	7	32	2	9	0	0	0	0
90	6	20	12	60	5	25	1	5	1	5	1	5
91	6	35	32	91	3	8	0	0	0	0	0	0

## Teaching to Objective:

The teacher discusses the objectives and expectations of the course.

89	19	22	5	22	7	32	4	18	1	5	2	9
90	19	20	14	70	5	25	1	5	0	0	0	0
91	19	35	30	86	5	14	0	0	0	0	0	0

At the beginning of the class, the teacher tells me what activities we will be doing during the class.

89	21	22	6	27	8	36	4	18	3	13	1	5
90	21	20	16	80	2	10	1	5	1	5	0	0
91	21	35	31	89	4	11	0	0	0	0	0	0

## Check for Understanding:

The teacher encourages questions if the material/lesson is not understood.

89	25	22	2	9	2	9	4	18	2	9	15	55
90	25	20	17	85	2	10	1	5	0	0	0	0
91	25	35	31	89	3	9	1	3	0	0	0	0

The teacher gets responses from the students to determine if they are learning during the lesson.

89	26	22	5	23	10	46	4	18	3	13	0	0
90	26	20	17	75	4	20	0	0	0	0	0	0
91	26	35	29	83	5	14	1	3	0	0	0	0

## Modeling:

The teacher's teaching methods encourage me to maintain interest during class.

90	28	22	9	41	7	32	0	0	0	0	0	0
91	28	20	17	85	2	10	2	10	1	5	0	0
92	28	35	30	86	5	14	0	0	0	0	0	0

The teacher explains difficult ideas in more than one way.

90	29	22	13	60	8	36	2	9	0	0	0	0
91	29	20	14	70	4	20	2	10	0	0	0	0
92	29	35	29	83	5	14	1	3	0	0	0	0



At the beginning of the class, the teacher tells me what activities we will be doing during the class.

90	21	27	3	11	4	15	8	30	5	19	7	26
91	21	48	26	54	13	27	5	10	3	6	1	2
92	21	33	23	69	5	15	0	0	1	3	4	12

Check for Understanding:

The teacher encourages questions if the material/lesson is not understood.

90	25	27	7	26	5	18	8	30	2	7	5	19
91	25	48	22	46	16	33	6	13	3	6	2	4
92	25	33	26	79	3	9	2	6	0	0	2	6

The teacher gets responses from the students to determine if they are learning during the lesson.

90	26	27	3	11	6	22	10	37	5	19	3	11
91	26	48	17	35	18	38	7	15	3	6	3	6
92	26	33	14	42	4	12	5	15	1	3	9	27

Modeling:

The teacher's teaching methods encourage me to maintain interest during class.

90	28	27	1	4	0	0	1	4	4	19	21	78
91	28	48	16	33	17	35	7	15	5	10	2	4
92	28	33	26	79	3	9	0	0	1	3	5	15

The teacher explains difficult ideas in more than one way.

90	29	27	2	8	3	11	3	11	5	18	14	52
91	29	48	19	48	13	27	10	21	3	6	2	4
92	29	33	22	67	5	15	3	9	0	0	2	6

Teacher Jo Reese

Yr	Q#	Rsp	A	A%	B	B%	C	C%	D	D%	E	E%
----	----	-----	---	----	---	----	---	----	---	----	---	----

Guided Practice:

I get enough practice with a skill or area before the class moves on to a new skill or area.

90	2	26	15	58	7	27	3	11	0	0	1	4
91	2	47	30	64	13	28	2	4	0	0	0	0
92	2	59	27	46	21	36	8	14	2	3	1	2

If I don't understand, the teacher finds a way to explain so that I do understand.

90	3	26	13	46	6	23	6	23	1	4	1	4
91	3	47	32	68	13	28	2	4	0	0	0	0
92	3	59	32	54	14	24	9	15	2	3	1	2

#### Independent Practice:

The teacher tests often enough to let me know how well I understand the material.

90	5	26	4	16	13	50	8	31	0	0	1	4
91	5	47	27	58	17	36	3	6	0	0	0	0
92	5	59	40	68	15	26	1	2	0	0	3	5

The teacher uses different ways (i.e., homework, questions, projects, written reports) to measure how much I know.

90	6	26	13	50	7	27	5	19	1	4	0	0
91	6	47	39	83	7	15	0	0	1	2	0	0
92	6	59	29	49	14	24	10	17	4	7	2	4

#### Teaching to Objective:

The teacher discusses objectives and expectations of the course.

90	19	26	12	46	6	23	6	23	1	4	1	4
91	19	47	25	53	15	32	6	13	0	0	1	2
92	19	59	32	54	15	26	9	15	3	5	1	2

At the beginning of the class, the teacher tells me what activities we will be doing during the class.

90	21	26	9	35	10	38	5	19	0	0	2	8
91	21	47	22	47	13	28	9	19	2	4	0	0
92	21	59	26	44	16	27	10	17	5	9	2	3

## Check for Understanding:

The teacher encourages questions if the material/lesson is not understood.

90	25	26	15	58	7	27	3	12	0	0	1	4
91	25	47	31	66	10	21	4	9	1	2	0	0
92	25	59	40	68	7	12	7	12	2	3	2	3

The teacher gets responses from the students to determine if they are learning during the lesson.

90	26	26	15	58	6	23	5	19	0	0	0	0
91	26	47	24	51	15	32	5	11	2	4	0	0
92	26	59	26	44	16	27	10	17	5	9	2	3

## Modeling:

The teacher's teaching methods encourage me to maintain interest during class.

90	28	26	10	38	7	27	7	27	1	4	1	4
91	28	47	26	53	12	26	7	15	2	0	0	0
92	28	59	38	64	11	19	6	10	0	0	3	5

The teacher explains difficult ideas in more than one way.

90	29	26	8	31	5	10	11	42	2	8	0	0
91	29	47	26	55	12	26	6	13	2	4	0	0
92	29	59	26	44	21	36	7	12	1	2	3	5

Teacher Dave Meredith

Yr	Q#	Rsp	A	A%	B	B%	C	C%	D	D%	E	E%
----	----	-----	---	----	---	----	---	----	---	----	---	----

## Guided Practice:

I get enough practice with a skill or area before the class moves on to a new skill or area.

90	2	46	24	52	14	30	3	6	4	9	1	2
91	2	29	19	66	7	24	0	0	0	0	3	10
92	2	34	33	97	1	3	0	0	0	0	0	0

If I don't understand, the teacher finds a way to explain so that I do understand.

90	3	46	22	48	3	7	8	17	5	11	8	17
91	3	29	22	76	4	14	0	0	0	0	0	0
92	3	34	33	97	1	3	0	0	0	0	0	0

#### Independent Practice:

The teacher tests often enough to let me know how well I understand the material.

90	5	46	11	24	9	20	13	28	12	26	1	2
91	5	29	20	69	4	14	3	10	1	3	1	3
92	5	34	30	88	3	9	0	0	1	3	0	0

The teacher uses different ways (i.e., homework, questions, projects, written reports) to measure how much I know.

90	6	46	7	15	2	4	4	9	5	11	28	61
91	6	29	18	62	7	24	1	3	1	3	3	10
92	6	34	28	82	3	9	1	3	2	6	0	0

#### Teaching to Objective:

The teacher discusses objectives and expectations of the course.

90	19	46	22	48	3	7	8	17	5	11	8	17
91	19	29	14	43	10	35	3	10	0	0	0	0
92	19	34	34	100	0	0	0	0	0	0	0	0

At the beginning of the class, the teacher tells me what activities we will be doing during the class.

90	21	46	17	37	4	9	8	17	4	9	13	28
91	21	29	23	79	2	7	1	3	1	3	0	0
92	21	34	30	88	4	12	0	0	0	0	0	0

#### Check for Understanding:

The teacher encourages questions if the material/lesson is not understood.

90	25	46	26	57	12	26	4	9	3	7	1	2
91	25	29	19	66	6	21	0	0	0	0	1	3
92	25	34	33	97	1	3	0	0	0	0	0	0

The teacher gets responses from the students to determine if they are learning during the lesson.

90	26	46	27	59	11	24	5	11	4	9	1	2
91	26	29	19	66	4	14	1	3	2	7	0	0
92	26	34	32	94	2	6	0	0	0	0	0	0

### Modeling:

The teacher's teaching methods encourage me to maintain interest during class.

90	28	46	16	35	2	4	9	20	3	7	16	35
91	28	29	18	62	5	17	2	7	0	0	1	3
92	28	34	32	94	2	6	0	0	0	0	0	0

The teacher explains difficult ideas in more than one way.

90	29	46	4	9	4	9	9	20	4	9	25	54
91	29	29	19	63	4	14	2	7	0	0	1	3
92	29	34	32	94	2	6	0	0	0	0	0	0

### Teacher Karen Holt

YR	Q#	Rsp	A	A%	B	B%	C	C%	D	D%	E	E%
----	----	-----	---	----	---	----	---	----	---	----	---	----

### Guided Practice:

I get enough practice with a skill or area before the class moves on to a new skill or area.

90	2	47	28	60	8	17	7	15	2	4	2	4
91	2	34	16	47	14	41	4	12	0	0	0	0
92	2	40	26	65	12	30	2	5	0	0	0	0

If I don't understand, the teacher finds a way to explain so that I do understand.

90	3	47	19	43	16	34	6	13	5	11	1	2
91	3	34	21	62	13	38	0	0	0	0	1	3
92	3	40	29	73	8	20	3	8	0	0	0	0



## Independent Practice:

The teacher tests often enough to let me know how well I understand the material.

90	5	47	10	21	21	45	11	23	4	9	1	2
91	5	34	15	44	12	35	3	9	3	9	1	3
92	5	40	23	58	9	23	7	18	1	3	0	0

The teacher uses different ways (i.e., homework, questions, projects, written reports) to measure how much I know.

90	6	47	19	40	17	36	6	13	3	6	2	4
91	6	34	25	73	8	24	0	0	1	3	0	0
92	6	40	31	78	7	18	2	5	0	0	0	0

## Teaching to the Objective:

The teacher discusses objectives and expectations of the course.

90	19	47	19	40	16	34	6	13	5	11	1	2
91	19	34	21	62	10	29	3	9	0	0	0	0
92	19	40	25	63	13	33	3	8	0	0	0	0

At the beginning of the class, the teacher tells me what activities we will be doing during the class.

90	21	47	9	19	12	26	12	26	4	9	10	21
91	21	34	26	72	17	21	1	3	0	0	0	0
92	21	40	33	83	6	15	2	5	0	0	0	0

## Check for Understanding:

The teacher encourages questions if the material/lesson is not understood.

90	25	47	26	55	11	23	7	15	2	4	1	2
91	25	34	15	44	16	47	3	9	0	0	0	0
92	25	40	28	70	9	23	4	10	0	0	0	0

The teacher gets responses from the students to determine if they are learning during the lesson.

90	26	47	28	61	8	17	7	15	2	4	2	4
91	26	34	16	41	13	38	4	12	1	3	0	0
92	26	40	26	65	11	28	4	10	0	0	0	0

## Modeling:

The teacher's teaching methods encourage me to maintain interest during class.

90	28	47	39	83	5	11	2	4	0	0	1	2
91	28	34	15	44	11	32	6	18	2	6	0	0
92	28	40	22	55	12	30	7	18	0	0	0	0

The teacher explains difficult ideas in more than one way.

90	29	47	2	4	3	6	10	21	16	34	16	34
91	29	34	15	44	13	38	5	15	1	3	0	0
92	29	40	23	58	11	28	5	13	2	5	0	0

APPENDIX C  
INTERVIEW QUESTIONNAIRE

Name:

I want to talk with you about the student rating of instruction instrument and procedure that we do each year.

1. Before we had this student rating of instruction, did you ever do a survey of the students to get an idea of how the class went or about your teaching procedures? Tell me about what you did and found.

2. What are your thoughts about the current system?

- A. What about the time of year the instrument is given?

B. Are the students surveyed frequently enough?

C. What do you think of the feedback you receive?

Do you feel it is generally an accurate picture of your teaching or how the class went?

D. What is your opinion of the students ability to make judgements about teaching? How much stock should be put in their responses?

E. Do you feel they would be better judges after they were away from the course or school for a while?

3. Tell me how you decided upon which classes to give the survey to.

A. Did the size of the class matter? In what ways?

B. Was the class elective or required? Do you think that mattered?

C. Do you feel the time of day the class is offered makes a difference? Why?

D. Did the level of the students make a difference when you were deciding upon what classes? How so?

E. Did you find any connection between the grade a student receives and the rating they give?

4. After you receive feedback from the students, tell me about how you interpreted the results?

A. Did you ever do anything differently in class because of the information you received? What?

B. Do you remember anything significant about the feedback you received?

C. In your opinion, do you think ratings are helpful in finding ways to improve instruction? How so?

5. How do you think student ratings should be used?

A. What part should they play in evaluation of a teacher's teaching?



6. What do you think are the strengths of the current system?
7. What are the weaknesses of the current system?
8. If you could make changes or adjustments to the current system in any way, what would they be?

9. Ten items in the instrument are based on effective components of instruction (guided practice, independent practice, checking for understanding, teaching to the objective, and modeling). The student ratings for you in these categories seemed to go up each year over the three-year period.

A. Can you remember if you did something in any of these areas to cause this to happen?

B. Do you recall if the ratings were helpful or informative?

C. Do you have any comments about the results as you see them over the three-year period?

10. Do you have any final comments about student ratings or about the procedure or instrument we currently use?